Fidelity Review: A scoping review of the methods used to evaluate treatment fidelity in behavioural change interventions.

O'Shea, O., McCormack, R., Bradley, J. M., & O'Neill, B. (2016). Fidelity Review: A scoping review of the methods used to evaluate treatment fidelity in behavioural change interventions. *Physical Therapy Reviews*, *21*(3-6), 207-214. https://doi.org/10.1080/10833196.2016.1261237

Published in:
Physical Therapy Reviews

Document Version:
Peer reviewed version

Fidelity review: A scoping review of the methods used to evaluate treatment fidelity in behavioural change interventions.

Orlagh O'Shea [1], Rosemary McCormick [1], Judy M. Bradley[2] Brenda O'Neill [1],

[1]Centre for Health and Rehabilitation Technologies, Institute for Nursing and Health Research, Ulster University, Northern Ireland. [2]NI Clinical Research Facility, School of Medicine, Dentistry and Biomedical Sciences, The Queen's University of Belfast, Belfast, Northern Ireland

Correspondence to: Dr Brenda O'Neill, [1]Centre for Health and Rehabilitation Technologies, Institute for Nursing and Health Research, Ulster University, Northern Ireland.
b.oneill@ulster.ac.uk

Objectives: To identify the definitions used for treatment fidelity in the behaviour change literature and to explore the extent to which the assessment of fidelity has been reported according to the five domains by Bellg et al.

Methods: Three data bases (Scopus, Medline Ovid and CINAHL) were searched. Results were limited to studies published between 2012 and 2015. Definitions/summaries of treatment fidelity used were recorded. Methods for assessing/monitoring treatment fidelity were extracted, summarised and categorised according to the five domains.

Results: Sixty-five papers were included for analysis. A definition of treatment fidelity was provided by n=34 studies; n=9 defined fidelity according to Bellg *et al.* In the context of treatment fidelity n=9 (13.8%) reported on study design; n=22 (33.8%) reported on an element of training of providers; n=59 (90.7%) papers reported on delivery of treatment; n=13 (20%) reported on receipt of treatment; and n=10 (15.3%) reported on enactment of treatment skills.

Conclusion: The definitions of treatment fidelity in the literature and the extent to which it has been reported were limited. Delivery of treatment was the most frequently reported component of treatment fidelity but other important aspects were poorly reported. The potential consequence of this is that translation of research interventions into clinical practice may not be optimised.

Key words: Treatment fidelity; behaviour change; physiotherapy; physical activity; exercise

Introduction

The concept of treatment fidelity has evolved over time;[1]and there does not appear to be one single agreed definition. Treatment fidelity can refer to all the mechanisms that ensure an intervention tests its hypothesis and that all the components of the intervention are implemented as outlined in the protocol. There does however appear to be an agreement in the literature of the importance of assessing and monitoring treatment fidelity. Firstly treatment fidelity increases the internal validity of a trial such that the results of the trial are directly attributable to the intervention.[2] Good treatment fidelity also increases the reproducibility of the trial by enhancing external validity; this increases to the extent to which the results can be generalised to the clinical setting. [1-3] Additionally optimisation of fidelity can also increase the statistical power of an intervention as the varability in delivery has been minimised. [1, 3-4] If the results of a trial are found to be non-significant and fidelity has not been monitored, it would be unclear if the results were due to an ineffective intervention or whether key elements of the trial were not implemented as planned. Conversely lack of attention to treatment fidelity could find an intervention to be effective due to extra treatment factors, potentially resulting in an ineffective intervention being found to be significant in a trial and subsequent implemented in clinical practice.[2-3, 5] Finally fidelity monitoring can aid researchers in moving forward and refining interventions, as it provides information on what components of the intervention were effective. [2S]

Treatment fidelity is of particular relevance to behavioural change interventions due to the complexity involved in targeting specific health behaviours for example physical activity. [1, 9,6]Due to the inherent nature of these complex interventions, there is greater capacity for variation especially when different research sites and treatment providers are involved. [3S] A review of behavoural change interventions between 1990-2000 found that 54% of studies did not report on intervention fidelity.[7] In an effort to rectify this Bellg et al.

as part of the National Institure of Health (NIH) Behaviour Change Consortium (BCC) identified five comprehensive domains under which treatment fidelity can be assessed and monitored or enhanced (Table 1). (1) design of study, (2) training providers (3) delivery of treatment (4) reciept of treatment (5) enactment of treatment skills. [1]

Table 1 National Institure of Health (NIH) Behaviour Change Consortium (BCC). Domains of Treatment Fidelity. Bellg et al. [1]

In the last decade, since the publication of NIH BCC recommendations on treatment fidelity, some studies have used these recommendations and it appears to be a useful model for monitoring and enhancing treatment fidelity.[2S,15S,16S,27S,343,54S,65S,8-10.]

Many aspects of physiotherapy include complex interventions (behavioural change, physical activity and exercise interventions). In order to ensure optimal translation of research findings into physiotherapy practice, knowledge of the fidelity of the trials that provide the underpinning evidence is important. Therefore, the aim of this paper is to identify how fidelity is defined in the literature, and to explore the extent to which reported fidelity is assessed/monitored in the published evidence on behaviour change, physiotherapy, physical activity interventions and exercise therapy and how the methods employed in this literature map to the five domains of the NIH BCC.

Methods

The overall approach will adopt scoping review methodology and included a six step framework: (1) identifying the research question; (2) searching for relevant studies; (3) selecting studies; (4) charting the data; (5) collating and summarising our result; (6) Consulting with key stakeholders (not applicable to this study). [11,12]

Identifying the research question: The research question which informed this review was "what methods are reported (in literature relating to behaviour change interventions, physical activity, exercise, physiotherapy) to assess/monitor treatment fidelity?"

Searching for relevant studies: A specialised search strategy was developed in consultation with the librarian for the School of Health Sciences, Ulster University. Two reviewers (OO'S, RMcC) independently and systematically searched three key databases (Scopus, Medline (Ovid), and CINAHL). Search words included "fidelity" OR "treatment fidelity" AND "behavio* change;" AND "physiotherapy" OR "physical therapy;" AND "exercise therapy;" AND "physical activity interventions." Searches were restricted to those conducted in humans and published in the English language. The literature was probed in preparation for this review and as a large volume of literature was available it was decided in advance of the search to limit the inclusion criteria to studies published from 2012-2015.

Selecting studies: Titles and abstracts were screened independently to identify relevant studies where "fidelity" was used in the context of our review aims. Search results were combined and duplicates removed. Only studies that included a clear method of assessing fidelity were included for data extraction. Review articles, case studies and commentaries were excluded, but held for discussion purposes. Full paper copies were retrieved and divided between the two reviewers; for training and standardisation, five articles selected at random were exchanged between reviewers and reviewed to assess agreement about whether studies met the inclusion criteria.

Charting the data: The research team met regularly to agree and refine the data extraction table. Ultimately the aims and objectives of the papers, a definition or summary of fidelity (if present) and the methods used to assess/measure fidelity were extracted and tabulated by each reviewer. The characteristics (design, population and number of participants) of the studies were also charted.

Collating and summarising our results: The extracted methods used to assess/measure fidelity were summarised and then mapped to the five domains as set out by NIH BCC framework. Table 1: design of study, training providers, delivery of intervention, receipt of the intervention and enactment of intervention skills. At the end of this process the reviewers met to agree the classifications and finalise the data extraction table.

Results

There were 65 papers included in this scoping review. The search results are available in figure 1.

One hundred and thirty seven full text articles were retrieved; n=65 of these were included and the remaining 72 papers were excluded for the following reasons: n=31 did not report a clear method of how fidelity was monitored or assessed and therefore did not meet the inclusion criteria. A further n=34 were review papers, 5 were editorial/commentaries, 1 was an opinion piece and the remaining 1 was a cross sectional questionnaire study.

The results of the data extraction are summarised in Table 2. Further details of the characteristics of the included papers, the definitions of fidelity and methods used to assess/monitor fidelity can be found in the E-supplement.

Fidelity definition

Thirty four of the 65 (52.3%) papers gave a definition/short summary of fidelity and of these 23 indicated a reference source for their definition, 21 different authors were referenced for definitions. The definition proposed by Bellg et al. was the most commonly cited definition of fidelity, cited by 9 [1], 21 different authors were referenced for definitions Most of the definitions centered around delivering the intervention as planned; 20 [6S,8S-9S,12S,17S-19S,21S-22S,24S,27S-28,S30S,36S,38S-39S,47S,56S,59S,60S] explicity used “delivery” in their definition while a further 8 used similar language for example “followed as planned,”implemented as planned”

"provided as intended."[5S,16S,23S,31S,35S,42S,57S,65S] Other definitions stated that fidelity is an important component of "verifying a cause-effect relationship within complex interventions,"[7S] and Hildebrand et al. included treatment differentiation in their definition.[57S]

Strategies for assessing/monitoring treatment fidelity mapped to the NIH BCC domains

Of the 65 papers included in this review only 2/65 (3%) included an assessment of all five domains; 39/65 (60%) papers assessed fidelity under one domain, 12/65 (18.5%) included two domains, 9/65 (13.9%) papers assessed fidelity under three of the NIH BCC components, and 3/65 (4.6%) addressed four of the five domains.

1. *Study Design*

Nine studies considered study design in their assessment/monitoring of fidelity (Table 2). Four of these studies reported on the underpinning theory.[2S,3S,54S,65S] Seven papers included a prior information on the dose to be delivered, ensuring it was the same between conditions.[11S,15S-16S,30S,34S54S,61S] Two of the included studies trained more than one provider as a strategy to allow for any setbacks.[2S,15S] Beck et al. used a specifc study design to minimise contamination and all providers in this study remained blind to the intervention content during the control period. [2S] Further strategies used to enhance fidelity relating to the domain of study design were incorporated by Winnet et al., where by they ensured that they would have sufficient statistical power to detect treatment effects. [15S]

2. *Training of providers*

Twenty two papers reported on the training of intervention providers in their assessment of fidelity (Table 2). Strategies reported to enhance provider training included standardisation of training so as all providers received a similar number of sessions or were given standard

training manuals. [2S,15S,22S,34S,46S,61S,65S] Role play or practice delivering the intervention was part of the training in nine studies[2S,14S,22S,44S,46S,52S,54S,64S-65S]; provider competence and adherence to the intervention components were usually assessed during these sessions. In efforts to minimise drift, refresher training was provided by Winnett et al. and others supervised or reviewed audio/video of sessions throughout the intervention and gave the providers feedback based on this; [15S] in one case the sessions were evaluated and if provders scored below a certain level of fidelity they were given additional training. [44S] Other strategies used included: seeking feedback on the training from the providers, [15S] using the results of the assessment of delivery to inform future training [17S] and the trainer reported if they had delivered the training as intended. [33S]

3. Delivery of treatment

Fifty nine included papers reviewed included an assessment of delivery (Table 2). Thirty nine studies assessed delivery of the intervention either by direct observation or through an evaluation of an audio or visual recording[1S-2S,6S-8S,10S,13S,17S,19S,20S,22S,-28S,32S-36S,39S-41S,44S-47S,51S,55S-58S,61S-65S] The number of actual treatment sessions assessed ranged from 10-100%. The criteria used to evaluate treatment delivery varied and included both objective checklists and subjective measures to evaluate the delivery of the intervention. For example in one study the raters reported on their "overall impression" of how the intervention was delivered[40S] another report evaluated the provider's engagement with the participants and whether the session was delivered in "a constructive and empowering manor."[56S] Other strategies used in the assessment of delivery included an effort to assess/measure the dose delivered (n=8).[8S,12S,23S,25S,31S,38S,42S,59S] The use of materials such as manuals used to enhance or aid delivery was used by four reviewed papers.[10S,15S,16S,62S]

4. Receipt of treatment

Thirteen of the papers included in this review reported an assessment of receipt (Table 2). Strategies use to assess receipt varied between authors and included ensuring that participants had an understanding of the intervention[15S, 11S,21S,60S]. Two authors made resources available to the participant so as they could perform the intervention activities. Other strategies included using online tracking codes to assess if participants accessed and received the material; [60S] one protocol reported that receipt would be assessed through brief questionnaires [27S] and Robbins et al. reported that receipt was assessed via providers' logs and assessment of audiotapes. [65S]

5. Enactment of treatment skills

An assessment of enactment of treatment skills was included by 10 of the studies (Table 2). The performance of the intervention skills was observed in the real life setting by one study[5S]; similarly two other reports used direct observation to examine the degree to which interventional changes took place. [18S,53S] Faulkner *et al.* used an objective measurement to assess if the treatment was being enacted in real life settings. [54S] Follow up contact to assess the enactment of the treatment skills was reported by two studies. [21S,30S]

Discussion

This review identified the definitions used for treatment fidelity and explored the extent to which the 5 domains of treatment fidelity are reported in the literature, and detailed the strategies used to capture these five domains. The definition by Bellg et al. was the most commonly cited definition for treatment fidelity. Most of the definitions provided centred around delivery of the intervention. The overall reporting of treatment fidelity is poor; only 40% reported on more than two of the five components. Treatment delivery was the most

frequently reported domain and this has been similarly noted in other papers.[30] Study design was the most under reported domain of fidelity with only nine studies including this domain in their analysis. There was a wide variation in the strategies used to assess/monitor fidelity across all domains.

The definition by Bellg et al. was the most commonly cited definition of treatment fidelity in the reviewed articles.[1] This definition centres mainly around reliability and validity, referring to both the strategies used to monitor and enhance these and the practices to ensure that the research reliably and validly tests the intervention. All of the reasons outlining the importance of measuring treatment fidelity as detailed in the introduction are directly realated to reliablity and validity (both internal and external) and it is likely that this definition provided by Bellg et al. was developed bearing in mind the benefits of ensuring good treatment fidelity.[1] Borrelli et al. also draw on upon this definition[7] and was cited by two reviewed studies.[2S,6S] However many of the papers in this review simply deduced fidelity down to the delivery; ensuring an intervetnion was delivered as intended. This simplified definition and concept of treatment fidelity may have influenced the methods used to assess treatment fidelity. This is evidenced through the results as treatment delivery was the most frequently assessed domain. The definition developed by Bellg et al. was developed by an expert group and we would encourage the use of this definition to aid in the standardisation of the assessment of treatment fidelity.

As treatment delivery was the most frequently reported domain it appears that authors have a good awareness of the importance of this. However all five components of fidelity are mutually exclusive; lack of consideration to any one category could potentially compromise the validity of the study. [7] For example if an intervention is found to be ineffective and the only domain of fidelity assessed was delivery which was high, it is possible that neglect of other domains may have caused the insignificant results; the providers may not have been

adequately trained or the study design may not have tested the hypothesis. There is some debate around the importance and relevance of all five domains. This review found enactment to be comparatively less well reported than the other four domains. Gearing et al. have conceptualised a treatment fidelity framework that does not include enactment as a core component of fidelity. [29] Gearing et al. also argue that enactment is a component of treatment efficacy rather than treatment fidelity; participants in a study may remain unwilling or unable to apply the treatment skills in real life settings despite the provider delivering the intervention as per protocol. [29] This is of particular importance to behavioural change interventions. The ultimate goal of behavioural change interventions is to change the participant's behaviour to enable them to engage with or carry out the treatment skills; if the participant remains unwilling to do so despite full consideration to the other four domains, perhaps this could then indicate that the treatment was ineffective. [2S] However, further work is required to wholly explore and agree this issue and come to a definitive conclusion on the relative importance of each of these five domains.

Study design was the most under reported component of fidelity and may have been over looked as an element of fidelity. Study design is an integral part of any intervention and impacts greatly on the ability of intervention to evaluate the hypothesis. [1] Only a small number of the studies in this review included a measure or assessment of study design when reporting fidelity. Bellg et al. outline specific criteria around study design so that the study can adequately test its hypothesis in relation to its underlying theory. [1] The theory which underpins interventions for behaviour change is important when designing an intervention, as it can provide a more in depth understanding of the processes of how the intervention might work[30], yet only four papers referred to a theoretical framework when reporting their fidelity assessment. Other reviews in various populations have found the reporting of the use of theories to underpin interventions ranged from 12-72%. [31-36] The aim of this review was to

summarise reported methods used to assess and monitor treatment fidelity; the evaluation of the study design was beyond the scope of this review and it is possible that papers reviewed included components of study design elsewhere.

This review focused on reports published since 2012. In 2011 Borrelli et al. published a checklist which further developed the NIH BCC framework into a 40- item checklist which was designed to assess the treatment fidelity of a study across all these five domains. [36] Despite the publication of the checklist preceding the publication of all the papers included in this review, it was only used by two of the studies [2S,15S] reviewed to help inform their assessment of treatment fidelity. Both these papers reported a comprehensive fidelity assessment; Beck et al. [2S] included four out of the five domains and Winnett et al. [15S] included all five domains. The lack of reporting of treatment fidelity in this review demonstrates the need for the use of a standard process or checklist to be used by authors so that none of the five components are overlooked. This checklist provides authors with a structured framework for which to monitor and assess all elements and components of treatment fidelity

Established reporting guidelines exist for the reporting and publication of clinical trials (CONSORT and TREND)[37-38] and protocols (SPIRIT).[39] None of these guidelines provide any specific guidance for the assessment and reporting of treatment fidelity. Although some of the components on these checklists do overlap with the NIHBCC guidelines, for example intervention content and dose. More recently Hoffman et al. 2014 published the TIDieR checklist (Template for Intervention Description and Replication) with the aim to improve the completeness of reporting and replicability of interventions.[40] This 12-item checklist contains two items of treatment fidelity (11 and 12). These items are ambiguous and limited in their description stating that only if intervention fidelity was it should be described and if assessed the extent to which it was delivered as planned should be

reported. It is however encouraging that fidelity is being included in these new guidelines. The monitoring, assessment and reporting of treatment fidelity would greatly benefit from the development of more explicit and compulsory reporting guidelines in line with the NIH BCC guidelines.

The inattention to treatment fidelity reported in this review may be due in part to the additional resources required to assess treatment fidelity. Assessing and monitoring fidelity requires increased time, equipment and personnel. This increased burden may concern researchers and funding agencies; Bellg et al. argue that not devoting these resources to treatment fidelity may be more costly in the longer term. Including a plan to assess and monitor treatment fidelity in a study can enhance the translation into clinical settings and reduce the likelihood of ambiguous results.[1] The physiotherapy research community have a vested interest in minimising the chance research can't be replicated in clinical practice. Lawton et al.[9S] provide an example of the importance of monitoring treatment fidelity for reliable and valid results; the authors found that a worksite physical activity intervention delivered across five sites was only found to significantly increase physical activity levels in one site where it was delivered with high fidelity.

Limitations

The actual documentation and reporting of fidelity within published papers is a central limitation to this review. This may be due in part to limitations on word count for journal publication. One way to overcome this issue is to provide online supplements so that the scientific community can access any additional information about the methods for assessing and monitoring treatment fidelity.

Finally the mapping of the reported methods of fidelity to the domains of fidelity as set out by the NIH BCC was based on reviewers' judgement. This may have led to some

misclassification of methods however attempts were made to reduce this as classifications were agreed by the two reviewers and regular meetings were held with a more experienced researcher throughout the process who was consulted when any disparity arouse.

Conclusion

In this scoping review we identified that there remains an inconsistency and paucity across the literature for the defining and reporting of methods for treatment fidelity assessment and monitoring in complex interventions. We recommned that future researchers should use the definition proposed by Bellg et al.[1] A fidelity framework such as that published by Borrelli et al. will support the comprehensive consideration and reporting of treatment fidelity in future research activities.[20] The use of this checklist to embed fidelity into clinical trials will ultimately enhance the translation of research into practice.

Funding

PhD funding: Department of Employment and Learning, Northern Ireland

References

1 Bellg A, Borrelli B, Resnick B, Hecht J, Minicucci D, Ory M, *et al*. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. Health Psychol. 2004;23:443–51.

2 Moncher, FJ, Prinz, RJ. Treatment fidelity in outcome studies. Clinical Psychology Review. 1991;11:247−266.

3 Resnick B, Bellg A, Borrelli B, DeFrancesco C, Breger R, Hecht J, *et al*. Examples of implementation and evaluation of treatment fidelity in the BCC studies: where we are and where we need to go. Ann Behav Med. 2005;29(Supplement):46–54.

4. Spillane V, Byrne M, Byrne MC, Leathem C, O'Malley M, Cupples M. Monitoring treatment fidelity in a randomized controlled trial of a complex intervention. J Adv Nurs. 2007;60:343–52

5.Hengeller S.W, Melton G.B, Brondino M.J, Scherer D.G,Hanley J.H. Multisystemic therapy with violent and chronic juvenile offenders and their families: The role of

treatment fidelity in successful dissemination. Journal of Consulting and Clinical Psychology. 1997; 75:821-833.

6 Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). Health Technol Assess. 2015 Nov;19(99):**0-187**

7 Borrelli B, Sepinwall D, Ernst D, Bellg A, Czajkowski S, Breger R, *et al*. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. J Consult Clin Psychol. 2005;73:852–60.

8 Robb S, Burns D, Docherty S, Haase J. Ensuring treatment fidelity in a multi-site behavioral intervention study: implementing NIH behavior change consortium recommendations in the SMART trial. Psychooncology.2011;20:1193–201.

9 Radziewicz R, Rose J, Bowman K, Berila R, O'Toole E, Given B. Establishing treatment fidelity in a coping and communication support telephone intervention for aging patients with advanced cancer and their family caregivers. Cancer Nurs. 2009;32:193–202.

10 Resnick B, Inguito P, Orwig D, Yu Yahiro J, Hawkes W, Werner M, Zimmerman S, Magaziner J. Treatment Fidelity in Behaviour Change Research; a case example. Nursing Research. 2005;54(2):139-143

11 Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci. 2010;5:69.

12 Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8:19–32.

13 Hansen WB. Introduction to the special issue on adaptation and fidelity. HealthEduc. 2013; 113:260-6.

14 Allen JD, Laura AL, Emmons KM. Fidelity and its relationship to implementation effectiveness, adaptation, and dissemination, in Dissemination and implementation research in health: translating science to practice, Brownson RC, Colditz GA, Proctor EK, editors. New York: Oxford University Press; 2012;(7):281–304.

15 Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Health services Research - process evaluation in randomised controlled trials of complex interventions. British Medical Journal. 2006;332: 413–416.

16 Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. British Medical Journal. 2008;337:1655.

17 Calsyn RJ. A checklist for critiquing treatment fidelity studies. Ment Health Serv Res. 2000;2(2):107-113.

18 Wyatt G, Sikorskii A, Rahbar MH, Victorson D, Adams L. Intervention fidelity:aspects of complementary and alternative medicine research. Cancer Nur.2010;33(5):331-342.

19 Bruckenthal P, Broderick JE. Assessing treatment fidelity in pilot studies assist in designing clinical trials: an illustration from a nurse practitioner community based intervention for pain. Advances in Nursing Science. 2007;30:72–84.

20 Borrelli B. The Assessment, Monitoring, and Enhancement of Treatment Fidelity in Public Health Clinical Trials. J Public Health Dent. 2011;71 (Supplement 1): 1-21.

21 Linnan, L and Steckler, A. Process evaluation for public health interventions and research.1st ed. San Francisco: Lossey-Bass; 2002:1-23.

22 Baranowski T, Stables G. Process Evaluations of the 5 a day projects. Health Education and Behaviour. 2000;27(2):157-166.

23 Saunders RP, Evans MH, Joshi P. Developing a process-evaluation plan for assessing health promotion program implementation: a how to guide. Health Promotion Practice. 2005;6(2):134-147.

24 Glasgow RE. RE-AIMing research for application: ways to improve evidence for family medicine. Journal of the American Board of family Medicine. 2006;29(1):11-19.

25 Glasgow R, Vogt T, Boles S. Evaluating the public health impact of health promotion interventions. The RE-AIM framework. Americian Journal of Public Health. 1999; 89. 1322-1327.

26 Fraser MW. Intervention research: Developing social programs. Oxford, New York: Oxford University Press; 2009.

27 Perepletchikova F, Kazdin AE. Treatment integrity and therapeutic change: Issues and research recommendations. Clin Psychol Sci Prac. 2005;12(4):365–383.

28 Sharpless BA, Barber JP. A conceptual and empirical review of the meaning, measurement,development, and teaching of intervention competence in clinical psychology. Clin Psychol Rev.2009; 29(1):47–56.

29 Gearing R, El-Bassel N, Ghesquiere A, Baldwin S, Gillies J, Ngeow E. Major ingredients of fidelity: a review and scientific guide to improving quality of intervention research implementation. Clin Psychol Rev. 2011;31(1):79–88.

30 Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. Health Psychol Rev. 2014 Aug 8:1e22.

31 Keogh A, Tully MA, Matthews J, Hurley DA. A review of behaviour change theories and techniques used in group based self-management programmes for chronic low back pain and arthritis. Manual Therapy. 2015;(20):727-735.

32 Painter J, Borba C, Hynes M, Mays D, Glanz K. The use of theory in health behaviour research from 2000 to 2005: a systematic review. Ann Behav Med. 2008;35:358e62.

33 Davies P, Walker A, Grimshaw J. A systematic review of the use of theory in the design of guideline dissemination and implementation strategies and interpretation of the results of rigorous evaluations. Implement Sci. 2010 Feb 9;5:14.

34 Prestwich A, Whittington C, Rogers L, Sniehotta F, Dombrowski S, Michie S. Does theory influence the effectiveness of health behaviour interventions? Metaanalysis. Health Psych. 2013;33:465-74.

35 French D, Olander E, Chisholm A, Sharry JM. Which behaviour change techniques are most effective at increasing older adults' self-efficacy and physical activity behaviour? A systematic review. Ann Behav Med. 2014;48:225-34.

36 Richardson J, Loyola-Sanchez A, Sinclair S, Harris J, Letts L, MacIntyre M, *et al*. Self-management interventions for chronic disease: a systematic scoping review. Clin Rehabil. 2014;28:1067-77.

37 Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010;8:18.

38 Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement.Am J Public Health 2004;94:361-6.

39 http://www.spirit-statement.org/ accessed on the 19th September 2016

40 Hoffmann T, Glasziou P, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. British Med J 2014;348:g1687

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

Author and study design	Aims of objectives	Population	Intervention(n=) control (n=)	Definition/description of fidelity	Methods of assessing fidelity
Bailey and Blair 2015[1S] Design: A multiple-baseline design	To examine the feasibility and outcomes of implementing the family-centred prevent teach reinforce model by replicating Sears et al. in a new sample.	Children with developmental disorders	N=3 boys aged 5-7	No definition.	♦All sessions were audiotaped. ♦Implementation fidelity was assessed using a specific checklist; which focused on the number of steps which were correctly implemented.
Beck et. al 2015[2S] Design: Study protocol for a step wedged randomised control trial.	To describe the methodology for promoting and facilitating the evaluation of intervention fidelity in The EAT (Eating As Treatment) project.	Patients undergoing radiotherapy for head and neck cancer.	Not reported; recruitment on-going	Treatment fidelity encompasses strategies designed to monitor and enhance the reliability and validity of behavioural interventions. [7]	♦**Study design:** Stated the underpinning theories and how these impacted the active components and the overall design of the study. The exact dose could not be set out given the flexibility of the designed intervention; providers completed a log and audio recorded sessions to verify this. Strategies were used to minimise contamination between groups (keeping providers blind to the intervention content during the control period and told not to discuss details of their intervention beyond their site, any apparent contamination will be analysed from audio recordings), the study team also provided for possible setbacks by training more providers than necessary and tailoring the training content and schedule to suit the providers. ♦**Provider training**: training was standardised for all providers as it was conducted by the same trainers using the same powerpoints, role play and discussions were used to ensure that the training was suited to the individual needs. Skill acquisition was assessed by self report assessments done before and after training, role plays were also videoed to assess skill acquisition. Ongoing supervision and feedback was provided to measure competence in delivering the intervention. Any concerns regarding clinician delivery of the intervention are discussed with the research team and raised with the clinician. This on-going supervision helped minimise drift in provider skills in addition to summarising key concepts on supplementary resources to prompt integration of training concepts into clinical practice. Booster training sessions were also completed. ♦**Delivery:** Supervision was used to monitor delivery. All sessions were audiotaped and assessed using an

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

					intervention specific checklist and standardised checklists to assess delviery. 20% of audiotapes were randomly seleted for assessment by trainined raters Providers had to score a minimum level on these checklists, if there were any concerns regarding delivery they were raised with the study team. Questionniares were used to collect information about the proviers previous training and clincal experience to account for any difference in providers; other questionnaires were used to assess dietician and patient perception of therapeutic alliance and the providers interpersonal skills were also measured. ♦**Receipt:** The authors felt that it was difficult to adopt the concept of receipt for this particular intervention and their interpretation of receipt for this was to focus on the degree to which the intervention was delivered.
Casey et. al 2015[3S] Design: Multiple baseline single-subject design	To evaluate the effects of a highly structured therapeutic skating intervention on motor outcomes and functional capacity.	**B**oys with autism spectrum disorder aged 7 and 10 years	Intervention n=2	No definition	♦Recorded attendance at specific time points. ♦At a particular time point specific measurements were taken of the tasks to be completed in the two trials.
Chesworth et. al 2015[4S] Design: *A priori* method of assessing fidelity of a Cluster randomised feasibility trial.	To explore fidelity to treatment delivery of the ICONS (Identifying Continence OptioNs after Stroke) intervention.	Adults post stroke	Intervention (n=40). Control (n=31)	"…the methodological strategies used to monitor and enhance the reliability and validity of behavioural interventions…[and]…the methodological practices used to ensure that a research study reliably and validly tests a clinical intervention" [1]	♦Clinical logs completed by the providers regarding the delivery of the intervention were reviewed.
Fortington et. al 2014[5S] Design: Observational	To measure the quality of exercise performance by players in FootyFirst, a coach-led, lower-limb injury prevention program.	Australian football players	Observed n=70	The extent to which a programme is followed as prescribed and adaptation is the extent to which a program is changed after implementation in a real world setting. [13-14]	♦Players were observed carrying out the exercises by two raters using a specifically designed checklist. Only observations that the raters agreed on were used for analysis.
French et. al 2015[6S]	To evaluate the fidelity of the IMPLEMENT	General practitioners	Intervention (n=59)	Intervention fidelity refers to both the	♦All workshop sessions were audiotaped and transcribed. The audio tapes were coded according to

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

Design: Comparison of planned and actual and observed versus self-assessed BCTs during the intervention.	intervention ; an interactive face-to-face educational intervention to improve general practitioner (GP) management of back pain		Control (n=53)	methodological strategies used to monitor and enhance the reliability and validity of delivery of interventions, and the extent to which an intervention as delivered is faithful to the intervention as planned. [1,7]	the presence of behavioural change techniques (BCTs). To establish reliability one transcript was coded by two raters and an agreement of 80% for the presence of a BCT had to be reached. One of these raters then coded the remaining transcripts 10% of which were randomly checked.
Fulkerson et. al 2015[7S] Design: Randomised control trial	To describe weight-related outcomes of the Healthy Home Offerings via the mealtime environment Plus study; a trial to prevent excess weight gain among youth.	Families (8-12 year old children and their parents)	Intervention (n=81) Control (n=79)	No definition	♦Pre-selected sessions were observed and delivery assessed using a standardised checklist.
Hanbury et. al 2015[8S] Design: Assessment of fidelity of an educational workshop	To summarise the fidelity assessment of a workshop designed to increase the uptake of a primary care alcohol screening recommendation.	Healthcare practitioners (general practitioners (GP), nurses, specialist alcohol service workers, healthcare assistants, dentists, health trainers)	N=62 participants (n=32 GPs, n=11nurses , n=4 specialist alcohol service workers, n=4 healthcare assistants, n=2 dentists, n=9 health trainers)	How well the delivery and receipt of the intervention mirrors the plans of those who have developed it – the intervention's fidelity – is increasingly recognised as an important determinant of its effectiveness. (No reference)	♦Sessions were observed and delivery assessed using a specified fidelity checklist, which rated the providers' adherence to the protocol. The providers' presentations were also examined for adherence and their presentation skills also rated. ♦Participant feedback regarding the style of the providers' delivery and the quality of the intervention was obtained. ♦Exposure/dose was evaluated by examining the attendance records to assess the number of targeted health professionals attending and the number of practices with representation.
Lawton et. al 2015[9S] Design: Fidelity analysis of a large matched-pair cluster randomised controlled trial	To test whether the effectiveness of a worksite physical activity intervention delivered in five work organizations varied as a function of fidelity.	Employees from 5 organisations across the UK (local council, hospital, bus company, government organisation, university)	N=1260	It is now widely acknowledged that when testing complex interventions via randomized controlled trials, it is important to collect data about how the intervention is delivered in practice (fidelity) and whether this varies according to the context.[1,15-16]	♦ (1)Adherence: assessed the extent to which each of the facilitators had delivered each of the 9 components. ♦ (2) Quaility of delivery was assessed self-report: facilitators were asked a number of questions regarding their perceptions of the quaility of the delivery and facilitators also reported on the number of hours they spent implementing the intervention. ♦ (3) Exposure: participants had to indicate the extent to which they had received each of the 9 components (yes/no) ♦ (4) Responsiveness was measured by exploring participants' perceptions of usefulness of each of the components of the intervention.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

					♦ (5) Engagement: particpants were asked to indicate whether they had taken part in the team challenges. Scores across all 4 domains was used to evaluate fidelity.
Martin et. al 2015 [10S] Design: A quasi experimental, pretest/ posttest design was used	To develop a sustainable, skill-based training program to assist older adults with their medication management	Community-dwelling older adults.	N=198	No definition	♦Academic research staff assisted with the development of a programme manual ♦Academic research staff attended all initial sessions delivered at each site to assess fidelity to the programme and materials and provided. Feedback was also provided.
McNamara et. al 2015[11S] Design: A single-cohort intervention study	To determine intervention fidelity by pharmacists for behavioural components of a complex educational intervention for cardiovascular disease (CVD) prevention.	Patients without established CVD, taking anti-hypertensive or lipid lowering therapy aged 50-74.	N=70	Demonstrable intervention fidelity is an important component of verifying a cause–effect relationship within complex intervention studies.[16]	♦ (1)Process indicators examined the appropriatess and suitability of the structure (taken from provider documentation); retention of participants and time taken to deliver the intervention. ♦ (2)Prcoess indicators were used to determine the appropriatene targetting and delivery of the intervention; (i) recruitment of particpants with uncontrolled risk factors (baseline documentation). (ii) Recommendations of goals to address particpants risk factors (baseline documentation). (iii) Patient agreement to pursue recommendations of strategies (taken from provider documentation). (iv) Development of strategies to address risk factors/goals (taken from provider documentation). (v) Indentification of barriers and enablers to behaviour change inititaion and maintenance (taken from provider documentation). ♦Providers also documetned their percieved success of behaviour strategies. ♦Self assessed perceived competence by providers to deliver the intervention was documented. ♦Providers perceived need for further patient support at completion of the intervention was documented.
Pawar et. al 2015[12S] Design: Cluster randomised control trial	To examine the feasibility of delivering an intervention promoting tobacco use cessation among school teachers.	School teachers	N=72 schools (n=36 control and n=36 intervention)	The extent to which intervention was delivered as planned ('fidelity'). (No reference)	♦Points were awarded if an intervention component was implemented, therefore the higher the score obtained the higher the fidelity.
Pincus et. al 2015[13S] Design: Randomised controlled feasibility trial	To test the credibility and acceptability of offering contextual cognitive behavioural therapy	Avoidant low back pain patients	N=89 (n=45 intervention, n=44 control)	No definition	♦The delivery of CBBT was assessed from audiotapes using a structured coding format. ♦The fidelity of the physiotherapists was established through (1) Exit interviews with a sample of particpants

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

	(CCBT) to patients with high fear avoidance who had been referred to physiotherapy.				(2) observations of one sessions per site the research team (3) exploration of the physiotherapy self report of session rating forms which detailed the components covered in each session.
Williams et. al 2015[14S] Design: Cluster randomised control trial	To investigate the role of Theory Planned Behaviour variables in predicting intention and objective walking behaviour in a sedentary general practice (GP) population.	Patients of GP practices aged 16-65 with one/more chronic condition, which increasing physical activity (PA) would have a positive effect and were sedentary (not meeting PA guidelines)	N=315 (n=136 intervention and n=179 control)	No definition	♦Providers were observed delivering the intervention before the trial commenced and were required to reach a minimum level of compentence before delivering the intervention in the trial.
Winnett et. al 2015 [15S] Design: Randomised Controlled Trial	To assess the efficacy of theory-based maintenance approaches varying by dose for persistently performing resistance training (RT) with the hypothesis that a higher-dose social cognitive theory (SCT) approach would produce greater RT adherence than lower-dose Standard.	Older adults (50–69 years), with a BMI of 25–39.9 kg/m2, all fitting pre-diabetes criteria.	N=170 enrolled in the initial 3 month phase. After the 3-month phase (N=159) were randomized to one of two conditions: SCT (intervention; N=79), or Standard (control; N=80).	No definition	♦**Design**: (i) The study design was based on a theory.(ii) The dose was set out before the intervention commenced.(iii) Specification of provider credentials. (iv)Ensured they had sufficient power to detect treatment effects. (v) Wave system of recruitment to match personnel. ♦**Training:** (i) The certificates of providers were checked before training. (ii) All providers received standardised initial training. (iii) Providers were given manuals. (iv) On-going supervision and feedback. ♦**Delivery**: (i) The providers were given session scripts to follow prompts for which points in the session to emphasise. (ii) Post session checklists were completed (iii) Sessions were randomly checked by the research team. (iv) Participants anonymously rated provider technical and interpersonal skills. (v) Sessions were supervised to maintain enthusiasm. (vi) Contamination was limited by using separate manuals for each condition and assigning any individuals with links to different groups. (vii) Participants reported on unsupervised sessions and were given feedback depending on group allocation.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

					♦**Receipt**: (i) All participants received hands on training and feedback for 3 months during the intervention. (ii) All participants can perform each exercise with proper form, range of motion, and degree of effort at the end of the intervention period. (ii) All participants were provided with a manual and instructions for the maintenance phase. ♦**Enactment**: (i) Participants completed transition sessions for unsupervised training; by the end of the transition participants were able to plan and report workouts.
Wyatt et. al 2015 [16S] Design: Randomised Controlled Trial	To examine the components of intervention fidelity, as put forth by the Treatment Fidelity Workgroup of the Behaviour Change Consortium at the National Institutes for Health (NIH-BCC Workgroup), within an ongoing acupressure study of breast cancer survivors with persistent cancer-related fatigue.	Breast cancer survivors	N=183	Fidelity consists of the measures taken to assure that an intervention is carried out as prescribed by the intervention protocol.[9,17-18]	♦**Dose parameters** A clear description of the dose to be given was set out and described from the start. ♦**Training** (i) Providers were trained to train participants in self-delivery by a certified acupuncturist. (ii) Demonstrations were conducted and the participants had to reach >/95% on the Acupressure Fidelity Form. (iii) Providers also received refresher training at a predefined point. (iv) Participant training: The correct technique was demonstrated to the participants. (vi) Participants then carried out the acupuncture with feedback and had to reach >/95% on the on the Acupuncture Fidelity Form before completing training. (v) Participants were also given an instruction manual and DVD. ♦**Self-delivery:** (i) Participants had a 3 week follow up session after the initial training to evaluate their technique. (ii) Feedback was provided to the participants and participants were required again to meet >/95% on the Acupressure Fidelity Form. (iii) The participants logged their sessions throughout the intervention and are given contact information in case questions arise during the intervention period. ♦**Intervention receipt:** (i) Participant logs were examined to evaluate receipt. (ii) Attrition rates were also used to examine the number of participants who completed the entire protocol. ♦**Enactment:** (i) This is on-going and not reported.
Avery et al. 2014 [17S] Design: Protocol for an open	To conduct an open pilot study to establish the acceptability, feasibility	Adults diagnosed with non-insulin	N=200 (n=100; intervention and n=100;	With so few primary studies explicitly utilising treatment fidelity	♦Consultations were videotaped (20-40%) and review appointments to assess adherence to and appropriate use of components of the intervention using a

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

pilot study and external pilot randomised control trial	and fidelity of the multifaceted intervention movement as medicine for type 2 diabetes in the primary care setting.	dependent type 2 diabetes for a minimum of 2 years.	control)	strategies to monitor and improve training for care providers (where training is offered), or to monitor the delivery of interventions to patients in practice, it is difficult to establish whether the interventions are being delivered as intended. Therefore it becomes impossible to decipher whether reported outcomes are a function of the intervention or 'non-intervention' factors.[3]	specifically developed checklist. Efforts will be made to record an equal number of consultations at each intervention time point. ♦The results of assessment of the delivery will be used to inform future training.
Baquero et. al 2014[18S] Design: Process Evaluation of a Randomised Control Trial	To describe a comprehensive process evaluation of an efficacious store-based intervention that increased store customers' fruit and vegetable consumption.	Shops That Serve Latino Immigrants in North Carolina; target population the customers of the sops	Four small-medium tiendas (n=2 intervention and n=2 control)	Fidelity was defined as the extent to which each of the intervention activities were delivered as intended, including the integrity and quality of the Intervention implementation. (No reference)	♦Process evaluation approach: Feedback was received from the employees and managers regarding the training. ♦Measured the amount of time managers and employees spent in training. ♦There was an assessment of how the funding for structural changes was allocated and which structural changes took place. ♦Assessed the degree to which the marketing campaign took place/was implemented; food demonstrations took place as planned and print materials were distributed as planned
Bryant et. al 2014[19S] Design: Three arm randomised control trial	To describe the processes in training physical therapists: (1) to deliver a standardized pain coping skills treatment and (2) to evaluate the effectiveness of that training.	People over the age of 50 with knee osteoarthritis	N=222 (Strengthening exercise n=75, pain coping skills training (PCST) n=74, strengthening exercises and PCST n=73)	Treatment fidelity, a term that refers to the consistent and reliable delivery of interventions.[1]	♦The quality of delivery of the intervention was assessed against previously standardised criteria from audio recordings of sessions (randomly selected 10% of recordings from both groups). Three measures of session's quality were used: (1) Adherence to each specific element (2) Physical therapist competence (3) Evaluated for demonstrated used of therapeutic skills.
Dewing et al. 2014[20S] Design: Comparison post training to follow up (12	To determine the impact of refresher training and supervision on counsellors'	Lay counsellors carrying out function related to health care	N=39	No definition.	♦Audio recordings were taken from two time points (1) recording per provider at time point 1 and up to 3 at time point (2) and rated with a specifically developed coding sheet as to whether they adhered to the

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

months)	proficiency in the intervention				protocol and according to (a) the clarity with which the counsellor explained the scale to the patient and (b) whether the counsellor was specific about the behaviour that they were asking the patient to rate themselves on. ♦Researchers also judged the quality of action plans agreed upon according to whether they appeared to have the potential to address the patient's adherence barrier or not.
Dyas et. al 2014[21S] Design: Qualitative study embedded in a pilot cluster randomised control trial	To investigate treatment fidelity of an educational intervention delivered to general practice (GP) teams; designed to improve the primary care management of insomnia.	Patients suffering from insomnia and general practice teams (GPs and practice nurses)	10 participants (n=6 patients, n=4 practitioners)	Treatment fidelity has been defined as the degree to which a treatment or intervention is delivered to participants as intended. [19]	♦Short telephone interviews were conducted with patients and practitioners who participated in the intervention to explore any breaches in fidelity. The conditions that they wanted to explore were set out a priori: (i) adherence to the intervention (ii) Patient receipt and understanding of the intervention (iii) Patient enactment. ♦The interviews were analysed to identify barriers and facilitators to these components of intervention fidelity and to understand why breaches in fidelity occurred.
Hardeman et. al 2014 [22S] Design: Randomised controlled trial	To develop a reliable coding frame for recorded consultations, and to describe the delivery and receipt of intervention and standard care components to understand how the intervention might have worked.	Patients with type 2 diabetes	N=211 (n=126; intervention. N=85; control)	Trial evaluations rarely include an assessment of the extent to which interventions are delivered and received as planned (fidelity), to what extent they are adapted, and what this means for long-term implementation and impact in routine clinical practice.[1]	♦Training was standardised for all nurses delivering the intervention. ♦The providers practiced intervention techniques during training. ♦All consultations were audiotaped and assessed adherence to scripted protocol. ♦Feedback was provided to nurses following listening to the audiotapes.
Kulwa et. al 2014 [23S] Design: Study protocol of a cluster randomised controlled trial	To implement and evaluate the effectiveness of a nutrition education package in improving infant and young child feeding practices, dietary adequacy and growth	Infants aged 6 months and their parents	Not applicable: Study protocol	Assess whether the intervention activities are implemented as planned (i.e. fidelity). (No reference)	♦Activity logs: A record will be kept of the amount of sessions conducted (with participants, health care workers, families and nutrition counsellors) and materials distributed. ♦Supervisory reports: a review of the providers' workbooks will be conducted to evaluate completeness, validity of documented information, referrals, appointments kept or missed. ♦Registration forms will record the number of community based nutrition counsellors trained and the number of health facility staff sensitised.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

					♦Pre-post test scores will be used to assess skill acquisition of providers was assessed before and after training. ♦Evaluation forms: To evaluate the quality of the training sessions was evaluated ♦Structured observations: Providers' interpersonal skills during home visits, use of intervention material, problem solving and confidence will be assessed.
Lorencatto et. al 2014[24S] Design: Fidelity assessment of a Cross-sectional study	To evaluate the fidelity of telephone-delivered behavioural support from the UK's national quitline service, using coded component behaviour change techniques (BCT's).	Smokers seeking cessation advice	75 sessions were audio recorded	Fidelity refers to the extent to which core intervention components are delivered as intended distinguished from how components are delivered such as quality.[20]	♦Identified BCTs in the treatment manual. ♦Audio recorded sessions (75) and assessed if the BCTs specified in the treatment manual were delivered in practice
McKenzie et. al 2014 [25S] Design: Randomised feasibility trial	To examine (1) operational feasibility of the programme; (2) participants' views of the programme; and (3) speech intelligibility, communication effectiveness and tongue and lip movement at four points.	Patients at least 3 months post stroke with no co-existing neurological condition and having dysarthria, with articulatory imprecision.	N=39 (n = 20, control and n = 19 intervention).	No definition	♦Monitored sessions to assess if the delivery was consistent with the protocol in relation to time distribution within sessions, therapy materials, and appropriate inclusion of modelling, practice opportunities, feedback, reinforcement, verbal reward, review, response correction, encouragement, communication maximization strategies, and achievement of 80% threshold success on stimulus sets before progression.
Neilson et. al 2014[26S] Design: Qualitative design	To investigate physical therapists' experiences and perspectives of a cognitive-behavioural informed training and intervention process as part of a randomized controlled trial involving adults with knee osteoarthritis.	Physical therapists	Eight physical therapists trained to deliver the programme	No definition.	♦Initial training was followed by additional formal mentoring and instruction, role playing, and performance feedback from a psychologist at each trial site over the course of 3 to 6 months ♦Audiotapes of training were reviewed by a psychologist to assess if the physical therapist was competent in delivering the intervention. ♦Audiotapes of the PT- patient interaction were reviewed throughout the study and feedback was provided to the PT from a psychologist.
Presseau et. al 2014[27S] Design: Two-armed cluster randomised controlled trial	To conduct a cluster randomised controlled trial to evaluate the effectiveness and costs of a theory-based behaviour	GP's, practice nurses/nurse practitioners, and healthcare assistants	Not applicable: study protocol (will be conducted in 44 GP practices)	Investigate whether the intervention was delivered as designed. (No reference)	♦**Delivery:** (i) Provider's will complete questionnaire-based facilitator report of delivery completed after each session. (ii) Consultations will be audio recorded and analysed using a checklist of the behavioural change technique (BCTs) to be delivered at each consultation

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

	change intervention targeting general practitioners (GPs) and nurses, to support improvement in the provision of high-quality care for people with type 2 diabetes.	working in the study practices actively engaged in providing diabetes care.			and whether the duration of the BCT changes over the course of the delivery period and between facilitators. (iii) Post intervention feedback forms will be distributed post intervention. ♦**Receipt and enactment** will be assessed through brief questionnaires delivered with the post intervention process evaluation.
Robbins et. al 2014[28S] Design: Process evaluation of a pilot intervention	To evaluate the reach, dose and fidelity of Guys Only Activity for Life (GOAL), a physical activity intervention programme and motivational interviewing techniques for 6th and 7th grade boys.	6th and 7th grade boys (USA).	2 schools (n=1; intervention and n=1 control. N=30 boys from each school)	Quality of intervention delivery or the extent to which the intervention was implemented in the manner and spirit in which it was intended.[21]	♦Observed delivery of a physical activity intervention using a survey adapted from other studies to assess delivery of the use of strategies to motivate, encourage or support the boys to increase their moderate vigorous physical activity. This was scored on a 4 point likert scale. ♦Motivational interviewing sessions were audio recorded. Two researchers were trained to evaluate these recordings and the Motivational Interviewing Code 3.1.1 was used to determine adherence to motivational interviewing. To further evaluate the delivery of the motivational interviewing the degree to which they assessed adherence to the underlying theory was assessed using a 4 point likert scale.
Van Schijindel- Speet et. al 2014[29S] Design: Process Evaluation of a Randomised Control trial	To describe the results of the process evaluation of a physical activity (PA) programme for people with intellectual disabilities (ID).	Adults (age 44+) with intellectual disabilities.	Eighty-one participants and 65 controls (age 44+) with mild or moderate ID.	Fidelity-implementation of the intervention.[22-24]	♦PA instructors reported directly to the researcher if a PA programme session was cancelled.
Washington et. al 2014[30S] Desing: Cohort	To advance the discussion of treatment fidelity in social and behavioural intervention research by analysing fidelity in an intervention study conducted within participating long term care settings of the Collaborative Studies of Long-Term Care.	Family members of relative in nursing homes and residential care/assisted living settings and staff of these settings.	N=6 nursing homes and n= 18 residential care settings (intervention). Control (not applicable).	The extent to which an intervention is delivered as intended.[25]	♦Study designed so as participants would receive a full dose of the intervention by attending all workshops. ♦Reminders were sent for upcoming workshops to encourage attendance and attendance at each workshop was recorded. ♦Participants were given a certificate of achievement upon completion and staff were given continuing education credits. ♦All supplies were made available to participants to ensure they could successfully perform these activities. ♦Follow up contact was made by the interventionist to see if a service plan had been created and if it was being

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

					followed as planned.
Almas et. al 2013[31S] Design: Group non-randomised cluster trial	To determine the feasibility and effectiveness of recruiting and retaining female preadolescents aged 9–11 years to both study arms and of implementing a 20-week school-based physical activity programme with the intervention group (treatment fidelity).	Girls aged 9-11. In Karachi.	N=280 (n=131 intervention group and n=149; control group)	Treatment fidelity was defined as the proportion of planned physical activity sessions actually held in the intervention group out of those planned. (No reference)	♦Recorded the amount of sessions delivered and reasons why session weren't delivered.
Bach et al. 2013[32S] Design: Feasibility and acceptability cohort study	To determine the feasibility and acceptability to physical therapists and patients of a cognitive behavioural pain self-management programme.	Physical therapy cohort and pain patient cohort	N=31 physical therapists and n=21 patients.	No definition.	♦A portion of consultations were audiotaped and scored with a predefined checklist. Fifty per cent were scored independently by two raters and the remainder were scored by a single rater.
Barber et al. 2013[33S] Design: Protocol for a pilot cluster randomised controlled trial	To describe the protocol for PIP Pre-schoolers in the Playground; a pilot cluster randomised control trial (RCT) of an outdoor playground-based physical activity intervention for children aged 18 months to 4 years; to assess the feasibility of conducting a full scale cluster RCT.	Parents and their children aged 18 months to 4 years old	Not applicable: Study protocol	No definition.	♦At the end of each session the trainer will record whether the training was delivered as intended. The providers being trained will also complete a short evaluation form at the end of each session to ensure skill acquisition. ♦3 Sessions at each intervention site will be observed and scored with a standardised form. ♦At the end of each session the provider will complete a form reporting whether the session was provided, the number attending and the activities provided.
Benzo et. al 2013[34S] Design: Pilot testing of intervention	To develop and test an intervention that focused on patient engagement for behaviour change in important aspects of the daily life in severe chronic obstructive pulmonary disease patients that can have impact on their perception of health and	COPD patients hospitalised for exacerbation	N=11	No definition.	♦**Study design** (i) strategies were utilised to ensure the treatment dose was the same within condition. (ii) Training provided to deal with different types of patients equally. (iii) All sessions recorded, with external monitoring. (iv) Interventionist self-monitoring of treatment delivery each session ♦**Training** (i) Standardised training, both materials and personnel. (ii) Training used recorded session review and role-play to help account for patient differences and interventionist differences in implementation style. (iii)

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

	hospitalizations and that could be integrated with pulmonary rehabilitation.				Interventionists were scored with pilot patients using session checklist. (iv) Interventionists used self-assessment with checklists. (v) Feedback was provided from recorded intervention session with interventionist. (vi) Interventionists asked to identify desired training topics to assist with intervention skill acquisition. (vii) Regular booster training sessions were provided. (viii) Reviewed sessions where the interventionist or fidelity monitor identified the session deviated from protocol. (ix) Regular debriefing meetings were held and training was centred according to needs, background, and clinical experience of the clinicians. ♦**Delivery**: (i) Delivery was standardised as an intervention protocol was used to guide each session. (ii) Recorded sessions and assessed them with a behavioural checklist completed by the fidelity Monitor. (iii) Providers completed a self-assessment checklist following each session. (iv) Case conferences were held in which providers discussed cases and trainer reviews skills and strategies.
Bergstrom et. al 2013 [35S] Design: Cluster randomised controlled trial	To investigate the effectiveness of a novel and complex intervention to improve diet and physical activity, targeting both caregivers and residents, in community residences for people with intellectual disabilities (ID).	Adults with ID and their caregivers	N=172 (N=90 ; intervention and n=80; control)	Intervention fidelity, defined as the extent to which a programme adheres to its programme theory.[26]	♦Providers' activity at network meetings was recorded and they were assigned points based on this. ♦Measured number of sessions held for residents (participants) and assigned points as per same.
Branscum et. al 2013 [36S] Design: Process Evaluation of a Group randomized controlled design.	To report the results of a comprehensive process evaluation for the "Comics for Health" program, a childhood obesity prevention intervention implemented at 12 after-school programs.	Children and adolescents	N=71 (n=37; control group, n=34; intervention group)	The extent to which the intervention was delivered as planned. (No reference)	♦Intervention sessions were observed with a structured tally sheet (author has established the readability and validity before use); which included a list of major tasks the provider was to complete to assess if the intervention was delivered as intended the provider also completed a separate checklist for self-check.
Gabbay et. al 2013 [37S] Design: 2-year randomized	To determine if the addition of nurse case managers trained in	Adults aged 18-75 with type 2 diabetes who	N=545 (n=232;control and n=313;	No definition.	♦Sessions were audio recorded and evaluated using a reliable and validated tool Behaviour Change Counselling Index to evaluate the delivery of the

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

controlled pragmatic trial	motivational interviewing to usual care would result in improved outcomes over two years in patients with type 2 diabetes who are at high risk for cardiovascular complications.	were at high risk for complications.	intervention)		motivational interviewing. ♦Feedback was given regularly based on these evaluations but diminished as the providers became more proficient. ♦The providers and two investigators met to review study progress biweekly or more frequently if needed.
Goode et. al 2013 [38S] Design: Evaluation of intervention delivered in the context of a cluster randomised control trial	To highlight what is optimally involved on the part of researchers to drive and facilitate successful health behaviour intervention implementation and evaluation in dissemination contexts.	Patients with type 2 diabetes or hypertension	Not reported (implementatio n paper)	Intervention fidelity or the extent to which a program is delivered as intended, or adheres to essential elements of the original evidence-based intervention.[25]	♦All providers were trained ♦Developed manuals for the providers and participants ♦Number of calls completed ♦Duration of calls completed ♦Participant use of program materials and satisfaction.
Lorencatto et. al 2013[39S] Design: Observational study	To evaluate a method for assessing fidelity of behavioural support; assess fidelity of delivery in two English Stop-Smoking Services; and compare the extent of fidelity according to session types, duration, individual practitioners, and component behaviour change techniques (BCTs).	Smoking cessation	N=21 recordings	Fidelity of an intervention refers to the extent to which interventions are delivered as intended, with adherence to specifications in intervention manuals. [1,20]	♦A proportion of consultations were obtained audiotaped. ♦Treatment manuals were coded according to an established taxonomy of BCTs. ♦Transcripts of the audiotapes were then coded according to the BCTs as per the treatment manual to assess delivery of the intervention.
Mars et al 2013[40S] Design: Fidelity assessment of a two-arm randomised controlled trial intervention	To demonstrate development and testing of tools, procedures to monitor and assess the intervention integrity of a complex intervention for chronic pain.	Chronic musculoskeletal pain	N=703 (n=403 intervention; n=300 control)	Intervention fidelity is defined as the use of methodological strategies to monitor and enhance the reliability and validity of behavioural programmes.[1]	♦All courses were audiotaped and fidelity was assessed under 3 domains. (i) Adherence: a component specific measure was designed to assess the delivery of key elements as described in the intervention facilitator's manual. (ii) Competence: A generic competence measure was designed to determine the extent to which the providers created an environment in which participants could share their experiences and learn new skills. (iii) Overall impression: Another measure was designed to reflect the extent to which the aims and objectives of the component were achieved and how the material was received in the group.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

Pfeiffer et. al 2013[41S] Design: Study protocol for a two-year randomized control trial (nested cohort design)	To observe the effects of a multi-component intervention on physical activity, sedentary behaviour, and physical activity energy expenditure in 3-5 year-old children; identify factors that associate with change in those variables; and evaluate the process of implementing the multi-component intervention.	3-5 year old children	Not applicable: Study protocol	No definition.	♦Direct observations and ratings of PA opportunities provided by teachers and children's PA during those opportunities (OSRAC-P, observational system for recording physical activity in children- preschool version). ♦Teachers' self-reports of intervention completeness, fidelity measures; barriers to implementation and children's responsiveness to the intervention were obtained. ♦The site directors' self-reports of practices related to physical activity with interviews were obtained.
Poston et. al 2013[42S] Design: Pilot randomised control trial	To determine if a complex intervention in obese pregnant women leads to anticipated changes in diet and physical activity behaviours and to refine the intervention protocol through process evaluation of intervention fidelity.	Obese pregnant women	N=183 (intervention; n=94, control; n=89)	If each component of the complex intervention was provided as intended. (No reference)	♦Health trainers (providers) completed audio diaries (130 recordings) reflecting on the fidelity and feasibility of the intervention delivery. ♦Measured if the intervention package was delivered as intended i.e. all consultations. ♦Group size was recorded. .
Scobbie et. al 2013[43S] Design: Process Evaluation	To examine the implementation, acceptability and perceived benefits of a goal planning and action planning framework in one community rehabilitation team with people recovering from stroke.	Stroke patients and health professionals (physiotherapists, occupational therapists, dietician, nurse and speech and language)	N=8 patients N=8 health professionals (n=2 occupational therapists; n=2 physiotherapist; n=1 dietician; n=1 nurse and n=2 speech and language therapists.)	No definition	♦Provider case notes for participants were reviewed to assess if the intervention was implemented as planned.
Sears et. al 2013[44S] Design: Mulitple baseline design	To examine the feasibility and potential efficacy of adapting the prevent-teach-reinforce model for use with two families of young children with	Autism spectrum disorder	N=2 boys (4 and 6 years old) and their families	No definition	♦Implementation fidelity was calculated as percentage based on the total number of correct intervention steps implemented divided by the total number of steps that were applicable. ♦Parents delivering the intervention were trained on a 1:1 basis. They practiced implementing the steps until

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

	autism spectrum disorders.				they could implement them with 90% accuracy. If the implementation scores fell below 80% at any point then additional coaching sessions were given. ♦The researchers reviewed video recordings with the parents and provided feedback.
Seo et. al 2013[45S] Desig: Prospective longitudinal design	To evaluate if the HEROES Initiative; a school-based childhood obesity prevention program based on the U. S. Centers for Disease Control and Prevention coordinated school health approach was able to effectively increase physical activity among elementary and middle school students who were exposed to the program for 18 months and to determine student and school-level predictors of success.	4th–8th grades from elementary and middle schools in Southern Indiana.	N=1091 (intervention only)	No definition	♦Interviewed school wellness co-ordinators, principals and cafeteria managers (on two occasions). ♦Observed the school environment assessing 9 specific domains relating to the intervention. Scores were awarded based on this observation to assess whether the intervention was being delivered as intended.
Sternfield et. al 2013[46S] Design: Randomised controlled 3 by 2 factorial trial	To describe the rationale for the 3 by 2 study design, to discuss issues relevant to intervention-specific methodology and implementation, and to present data on recruitment, eligibility, and baseline characteristics	Post-menopausal women	N=355	No definition	♦Training was standardised and all providers were given a study manual. ♦During training mock yoga classes were conducted and all yoga instructors were given training CDs, DVDs and handbooks. ♦Exercise trainers were given detailed written instructions regarding prescription and progression of exercises. ♦The importance of strict adherence to the intervention protocol was emphasised repeatedly during trainings. ♦Fidelity of the yoga intervention was monitored through the completion of a form by an unblended staff member and the yoga instructors communicated weekly via email with the Seattle investigators to describe how classes were proceeding and if they had any questions or concerns. ♦Fidelity of the exercise intervention was monitored whereby one session a week was observed to ensure fidelity to the protocol using a quality control checklist.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

					The exercise trainers completed a log to ensure the prescribed dose was being achieved. Exercise trainers, supervisors and experts in exercise training had regular conference calls to resolve any issues. ♦For both exercise and yoga, a list of "Frequently Asked Questions" was compiled and distributed monthly to ensure a standardized approach to any issues that arose that had not been specified in the protocol. In addition, site visits were conducted.
Wilner et. al 2013 [47S]	To evaluate the impact of a staff-delivered manualised cognitive behaviour therapy anger management intervention on reported anger among people with mild to moderate intellectual disabilities, and anger coping skills, aggression, mental health, quality of life and costs of health and social care; factors that influence outcome; and the experience of service users, lay therapists and service managers.			Therefore, treatment integrity or fidelity checks are needed, in order to be able to monitor the extent to which treatments are delivered appropriately.[2]	♦Fidelity was monitored by direct observation. A pair of observers attended selected sessions to monitor fidelity. ♦An existing checklist (CTS-Psy66) was adapted to monitor the fidelity of the intervention. Additionally monitors made global ratings on a ten point's scale of fidelity to the manual, group process, principles of CBT and a single overall rating. Observers then compared their results and discussed any differences to come to a consensus decision.
Zheng et. al 2013[48S] Design: Randomised Control Trial	To design a system to support the fidelity of intervention delivery and efficient capture of qualitative and quantitative process data for a telephone-delivered behavioural change counselling intervention to increase physical activity and function after total knee replacement surgery.	Patients with advacned knee osteo arthritis post total knee replacement	Not reported	No definition	♦On screen documentation and prompts guided the providers through the consultation to deliver all components.
Bodde et. al 2012[49S]	To conduct a formative and process evaluation of	Adults with intellectual	N=21 (n=21 women and	No definition.	♦Providers were instructed to use an exact script. ♦On four random occasions the provider's adherence to

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

Design: Formative and process evaluation strategies	the Promoting Health through Physical Activity Knowledge and Skills curriculum which was designed to increase the physical activity knowledge and skills of adults with intellectual disabilities.	disabilities.	n=21 men)		the script was assessed.
Broekhuizen et. al 2012[50S] Design: Parallel randomised control trial	To evaluate the efficacy of an individualised tailored lifestyle intervention on physical activity, dietary intake, smoking and compliance on statin therapy in people with Familial Hypercholesterolemia	Adults with familial hypercholestero lemia	N=340 (n=181; intervention and n=159 control)	No definition.	It was assessed whether face-to-face counselling sessions were implemented as planned according to motivational interviewing (MI) guidelines (i.e. MI fidelity) was assessed by two MI experts, following the Motivational Interviewing Treatment Integrity code (MITI 3.1.1.)
Brookman-Frazee et. al 2012[51S] Design: Pilot single armed intervention	To examine the feasibility of training community mental health therapists to deliver a package of evidence-based practice strategies to children with autism spectrum disorders and challenging behaviours, and their parents with routine services.	Children with autism spectrum disorder and community based mental health therapists.	N=13 community based mental health therapists and n=13 children with ASD	No definition	Three methods were used to measure fidelity: ♦Treatment planning phase fidelity: treatment planning forms were reviewed by intervention developers to assess to adherence to key elements. ♦The active treatment phase session fidelity treatment: treatment sessions were observed. This included ratings on 3 required within sessions therapist behaviours. Each therapist behaviour had associated therapist strategies which guided a rating on a 4 point Likert scale. ♦ Therapists completed a web based survey after the training period. For each intervention, the step therapists rated the extent to which they completed each step.
Cate et. al 2012[52S] Design: Protocol for a randomised control trial	To determine whether additional education and advice about glaucoma using a Behaviour Change Counselling intervention, improves adherence with topical anti-glaucomatous therapy.	People with glaucoma	Not applicable: Study protocol	No definition	♦The providers information provision was assessed in terms of adherence to the BCC template and consultation style assessed using Behavioural Change Counselling Index via a video recorded session with an actor patient. The video recorded role-play session were independently reviewed according to the BBCI criteria by the Motivational Interviewing (MI) coach and two experts in MI independent to the research study. ♦Individualised written feedback was provided to the providers.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

Cowan and Devine 2012[53S] Design: Process evaluation of a quasi-experimental design	To evaluate the implementation of a controlled, 6 week environmental and educational intervention to improve dietary intake and body composition, and to study the association if implementation fidelity with diet and body composition outcomes.	Residents of drug treatment facilities	N=107	No definition	♦Food environment changes were assessed through direct observations of reviewed shopping lists, weekly menus and food inventories in each of the six facilities, and observed meals.
Faulkner et. al 2012 [54S] Design: Fidelity assessment of a feasibility intervention.	To describe the components of intervention fidelity, the complexity of measurement when conducting research with youth and families, and strategies for measuring intervention fidelity.	Adolescents with type one or type 2 diabetes.	N=50	Intervention fidelity refers to the methodological strategies used to monitor and enhance the reliability and validity of behavioural interventions. [1]	♦**Study Design:** The intervention was built on a strong theoretical foundation for exploring behaviour change with an evidence base to support it. Treatment dose and intervention length were set out from the start. ♦**Training of providers:** (i) A detailed study manual was developed. (ii) Providers learnt the study protocol and proper clinical etiquette for recruitment and professional communication with participants.(iii) Role play was also done so research assistants (RAs) could become more familiar with recruitment scripts, use of equipment and conducting home visits and fidelity checklists for the personalised exercise programme. ♦**Delivery:** (i) Fidelity checklists were completed at each home visit. (ii) The study team met weekly to discuss home visits fidelity checks, accelerometer downloads and any questions from the RAs could also be addressed. ♦**Receipt:** Feedback was obtained from the participants about refinement of the intervention to further enhance sustainability of exercises. ♦**Enactment:** Accelerometer recordings over the 16 weeks served as a measure of enactment.
Gallanter et. al 2012[55S] Design: Retrospective pre-post design	To further explore the effectiveness of in-home parent child interaction therapy with a diverse sample of parent–child dyads by using data from a child maltreatment	Families/parents at risk of maltreating children	N=83 clinical records of families were reviewed.	No definition	♦The supervisor monitored two sessions per year to ensure consistency with the protocol.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

	prevention program.				
Heideman et. al 2012 [56S] Design: Pilot study of single arm intervention	To assess the fidelity, feasibility and acceptability of a prevention program for overweight first degree relatives of type 2 diabetes patients intervention prior to starting the randomized controlled trial.	Individuals with a family history of type 2 diabetes.	N=21	Asses the fidelity (where intervention modules delivered as intended). (No reference)	♦All the sessions were observed and findings recorded on a specifically developed checklist based. Observers checked whether all modules were delivered and all objectives for participants were covered; observers reported on the engagement of participants by looking at interactions between trainer and participants and among participants; and observed whether the sessions were delivered in a constructive, empowering atmosphere.
Hildebrand et al 2012 [57S] Design: Fidelity assessment of randomised control trial	To describe the development of methods to train and supervise therapists to attain adequate treatment fidelity in a treatment development project involving a novel occupational therapy and physical therapy based intervention.	Older adults who are in short term skilled nursing facilities (SNF) following a disabling medical event	N=26 (n=14; intervention group, n=12 control)	Treatment fidelity comprises two key aspects: 1) treatment integrity, that is, demonstrating that therapists carry out the intervention with adequate levels of adherence and competence to the treatment model or protocol; and 2) treatment differentiation, that is, ensuring that the experimental intervention condition differs from a control condition (i.e., showing much higher adherence and competence to the treatment model.[27,28]	♦All sessions were videotaped and rated with a checklist specifically developed to rate treatment adherence and competence that quantified behaviours consistent with the intervention. Observations for fidelity ratings were done 12 months after therapists training while they were receiving on going supervision.
Hollands et. al 2012[58S] Design: Parallel group, cluster randomised controlled trial	To test the hypothesis that communicating risk of developing Crohn's disease based on genotype and that stopping smoking can reduce this risk motivates behaviour change among	Smokers who were first degree relatives of probands with Crohn's disease	N=497 (n=251; intervention; n=246 control)	No definition	♦Reviewed a random selection of audiotapes to assess fidelity to the protocol.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

	smokers at familial risk.				
Irvine et al 2012 [59S] Design: Process Evaluation of text message delivered intervention	To assesses the utility of novel techniques for process evaluation involving no face to face contact.	Men aged 25 to 44 years, who lived in areas of high social deprivation and had regular episodes of heavy drinking.	N=67 (n=34 ; Intervention n=33;control)	The fidelity of delivery of the intervention (the extent to which the text messages were delivered as intended). (No reference)	♦Recorded how many text messages were delivered.
Knowlden and Sharma 2012[60S] Design: A Feasibility and Efficacy Randomized Controlled Trial (protocol)	To evaluate the efficacy of the Enabling Mothers to Prevent Childhood Obesity Through Web-Based Education and Reciprocal Determinism program, an Internet-based, theory-driven intervention for preventing childhood overweight and obesity.	Mothers with children aged 4-6.	Not applicable: Study protocol	Implementation process evaluation is a specific type of process evaluation that examines fidelity of program delivery. Assessment of implementation allows the researchers to ensure the program was delivered to the participants in the prescribed fashion. Failure to evaluate program fidelity can make it difficult to confirm whether non-significant program outcomes were due to ineffective intervention components or inadequate transference of intervention deliverables. (No reference)	♦Log-in codes and tracking data will be used to assess whether the website and subsequent module materials were accessed. The date and duration of activity will be logged to assess whether audio-visuals were viewed and adequate time was spent to complete each activity. ♦Online, interactive worksheets and module quizzes will have forced-response validation to gauge transference of information. ♦Reminder emails will be sent to assess promotion. ♦At the completion of the intervention, respondents will be requested to complete an open-ended questionnaire regarding acceptability and perceived usefulness of the program. Additionally, data regarding maintenance of confidentiality will be collected.
Llewellyn et al 2012[61S] Design: Multicentre randomised control trial (protocol)	To examine the impact of motivational interviewing augmented with information provision and behavioural skills building,	Men who have sex with men (MSM) prescribe PEP for HIV	Not applicable: Study protocol.	Assessing the fidelity of the treatment is an important component of successful research dissemination. (No	♦Study design has ensured there will be the same dose between conditions. ♦Reduction of differences within treatments will be ensured by the use of one trained interventionist. ♦Interventionist skill acquisition and minimising 'drift' in

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

	over and above usual care, on risky sexual behaviour in men who have sex with men prescribed post exposure prophylaxis (PEP) after potential sexual exposure. A secondary aim of this research is to examine the impact of the intervention on adherence to PEP.	following sexual exposure		reference.)	interventionist skills will be minimised by the development and use of a treatment manual with the provision of feedback. ♦Audiotape sessions and coded using a validated instrument to ensure delivery and provide feedback to the provider. ♦Provider to complete a checklist after each session to remind him to include appropriate skills and content. ♦An advisory board will be used to monitor whether treatment protocol has been adhered to during recruitment and intervention period.
McCurry et. al 2012[62S] Design: Pilot randomised control trial	To investigate the feasibility of implementing a Sleep Education Program (SEP) for improving sleep in an adult family home residents with dementia and the relative efficacy of SEP compared to usual care (control) in a pilot randomised control trial	Adult family home (AFH) caregivers and residents with dementia and sleep disturbances	N=84 (n=37 AFH caregivers; n=47 residents)	No definition	♦**Delivery**: Providers were given a written manual with materials for each session. A checklist was completed after each session indicating which treatment topics had been covered. All sessions audiotaped and reviewed by investigator who provided feedback re adherence to treatment protocol. ♦**Receipt:** Staff-caregiver attendance at the sessions and clinical impressions were rated by a trainer after each session. The trainer also recorded whether staff-caregivers were able to identify specific behaviours and develop plans based on these behaviours for the week. ♦**Enactment:** The trainer reviewed homework at every session, rated homework compliance and assisted staff-caregivers in problem-solving.
Moore et. al 2012[63S] Design: A Mixed methods study	To examine implementers views on delivering motivational interviewing (MI) within an exercise referral scheme and consistency of consultations with MI before and after a 2 day workshop.	Exercise professionals and area coordinators delivering the Welsh National Exercise Referral Scheme.	N=37 (n=27 exercise professionals and n=10 coordinators)	No definition	♦Recordings of consultations were assessed using Behaviour Change Counselling Index. ♦Coders then estimated whether professionals spoke for more than half, about half or less than half of the consultation time. ♦Pre training to fidelity MI was compared with post training fidelity.
Morganstrern et. al 2012[64S] Design: Pilot 3 armed intervention study	To test the causal role of key hypothesized active ingredients and mechanisms of change within motivational interviewing (MI) in reducing drinking.	Adults between 18-65 with alcohol use disorder	N=89 (N=29 motivational interviewing (MI); n=30 SOMI (Spirit Only MI); n=30 SC (Self Change)	No definition	♦**Training**: Videotapes of practice cases were reviewed to ensure fidelity to the protocol. Performance was then reviewed and therapists were required to meet a certain level of fidelity before treating participants. ♦**Delivery** (i) 30% percentage of sessions were observed and assessed for fidelity to MI using the MI integrity code 3.0 to assess fidelity from the observer

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

					perspective. (ii) The modified version of the therapy session report was used to assess for fidelity from the client perspective.
Robbins et. al 2012[65S] Design: Two-group pretest posttest quasi-experimental study	To describe the methodology and findings related to the treatment fidelity of face-to-face motivational interviewing sessions involving middle school girls and a school nurse to help the girls increase their moderate to vigorous physical activity.	Middle school girls (10-14 years)	N=37	Developing, implementing, and evaluating a treatment fidelity plan is a time-consuming, but important, process for researchers to ensure that an intervention has been implemented as intended and accurately tested [1]	♦**Study design:** The underlying theory is stated and how it was congruent with clinical process. ♦**Training:** An additional provider was trained to allow for potential setbacks. Training was standardised and the providers were given an intervention manual. The providers did role play and were given feedback as part of the training. ♦**Delivery and receipt:** The providers kept logs of the sessions. All sessions were audiotaped and some were randomly selected for assessment.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

References

1S. Bailey KM, Blair KC. Feasibility and potential efficacy of the family-centered Prevent-Teach-Reinforce model with families of children with developmental disorders. Res Dev Disabil 2015 12;47:218-233 16p.

2S.Beck AK, Baker A, Britton B, Wratten C, Bauer J, Wolfenden L, *et al*. Fidelity considerations in translational research: Eating As Treatment - a stepped wedge, randomised controlled trial of a dietitian delivered behaviour change counselling intervention for head and neck cancer patients undergoing radiotherapy. Trials. 2015;16(1).

3S. Casey AF, Quenneville-Himbeault G, Normore A, Davis H, Martell SG. A Therapeutic Skating Intervention for Children with Autism Spectrum Disorder. Pediatric Physical Therapy. 2015;27(2):170-177.

4S.Chesworth BM, Leathley MJ, Thomas LH, Sutton CJ, Forshaw D, Watkins CL. Assessing fidelity to treatment delivery in the ICONS (Identifying Continence OptioNs after Stroke) cluster randomised feasibility trial. BMC Med Res MethodoL. 2015;15(1).

5S. Fortington LV, Donaldson A, Lathlean T, Young WB, Gabbe BJ, Lloyd D, *et al*. When just doing it' is not enough: assessing the fidelity of player performance of an injury prevention exercise program. Journal of Science & Medicine in Sport. 2015 May;18(3):272-277.

6S.French SD, Green SE, Francis JJ, Buchbinder R, O'Connor DA, Grimshaw JM, et al. Evaluation of the fidelity of an interactive face-to-face educational intervention to improve general practitioner management of back pain. BMJ Open 2015;5(7).

7S.Fulkerson JA, Friend S, Flattum C, Horning M, Draxten M, Neumark-Sztainer D, *et al*. Promoting healthful family meals to prevent obesity: HOME Plus, a randomized controlled trial. Int J Behav Nutr Phys Act. 2015;12(1).

8S. Hanbury A, Farley K, Thompson C, Wilson PM. Assessment of fidelity in an educational workshop designed to increase the uptake of a primary care alcohol screening recommendation. J Eval Clin Pract. 2015;21(5):873-878.

9S. Lawton R, Mceachan R, Jackson C, West R, Conner M. Intervention fidelity and effectiveness of a UK worksite physical activity intervention funded by the Bupa Foundation, UK. Health Promot Internation. 2015;30(1):38-49.

10S. Martin BA, Chewning BA, Margolis AR, Wilson DA, Renken J. Med Wise: A theory based program to improve older adults' communication with pharmacists about their medicines. Res Soc Adm Pharm. 2015.

11S. McNamara KP, O'Reilly SL, George J, Peterson GM, Jackson SL, Duncan G, et al. Intervention fidelity for a complex behaviour change intervention in community pharmacy addressing cardiovascular disease risk. Health Educ Res. 2015;30(6):897-909.

12S. Pawar PS, Nagler EM, Gupta PC, Stoddard AM, Lando HA, Shulman L, *et al*. Tracking intervention delivery in the 'Tobacco-Free Teachers/Tobacco-Free Society' program, Bihar, India. Health Educ Res. 2015 Oct;30(5):731-741.

13S Pincus T, Anwar S, McCracken LM, McGregor A, Graham L, Collinson M, *et al*. Delivering an Optimised Behavioural Intervention (OBI) to people with low back pain with high psychological risk; results and lessons learnt from a feasibility randomised controlled trial of Contextual Cognitive Behavioural Therapy (CCBT) vs. Physiotherapy. BMC Musculoskelet Disord. 2015;16(1).

14S. Williams SL, Michie S, Dale J, Stallard N, French DP. The effects of a brief intervention to promote walking on Theory of Planned Behavior constructs: A cluster randomized controlled trial in general practice. Patient Educ Couns. 2015;98(5):651-659.

15S. Winett RA, Davy BM, Savla J, Marinik EL, Kelleher SA, Winett SG, *et al*. Theory-based approach for maintaining resistance training in older adults with prediabetes: adherence, barriers, self-regulation strategies, treatment fidelity, costs. Transl Behav Med.2015;5(2):149-159.

16S. Wyatt G, Frambes D, Harris RE, Arnedt JT, Murphy SL, Zick SM. Self-administered acupressure for persistent cancer-related fatigue: Fidelity considerations. Altern Ther Health Med. 2015;21(4):18-23.

17S Avery L, Sniehotta FF, Denton SJ, Steen N, McColl E, Taylor R, *et al*. Movement as Medicine for Type 2 Diabetes: Protocol for an open pilot study and external pilot clustered randomised controlled trial to assess acceptability, feasibility and fidelity of a multifaceted behavioural intervention targeting physical activity in primary care. Trials. 2014;15(1).

18S Baquero B, Linnan L, Laraia BA, Ayala GX. Process Evaluation of a Food Marketing and Environmental Change Intervention in Tiendas That Serve Latino Immigrants in North Carolina. Health Promot Pract. 2014;15(6):839-848.

19S. Bryant C, Lewis P, Bennell KL, Ahamed Y, Crough D, Jull GA, *et al*. Can physical therapists deliver a pain coping skills program? An examination of training processes and outcomes. Phys Ther.2014;94(10):1443-1454.

20S Dewing S, Mathews C, Cloete A, Schaay N, Simbayi L, Louw J. Lay counselors' ability to deliver counseling for behavior change. J Consult Clin Psychol. 2014;82(1):19-29.

21S Dyas JV, Togher F, Siriwardena AN. Intervention fidelity in primary care complex intervention trials: qualitative study using telephone interviews of patients and practitioners. Qual Prim Care. 2014 Feb;22(1):25-34.

22S Hardeman W, Lamming L, Kellar I, De Simoni A, Graffy J, Boase S, *et al*. Implementation of a nurse-led behaviour change intervention to support medication taking in type 2 diabetes: Beyond hypothesised active ingredients (SAMS Consultation Study). Implement Sci. 2014;9(1).

23S. Kulwa KB, Verstraeten R, Bouckaert KP, Mamiro PS, Kolsteren PW, Lachat C. Effectiveness of a nutrition education package in improving feeding practices, dietary adequacy and growth of infants and young children in rural Tanzania: rationale, design and methods of a cluster randomised trial. BMC Public Health. 2014;14:1077.

24S Lorencatto F, West R, Bruguera C, Michie S. A method for assessing fidelity of delivery of telephone behavioral support for smoking cessation. J Consult Clin Psychol. 2014;82(3):482-491.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

25S. Mackenzie C, Muir M, Allen C, Jensen A. Non-speech oro-motor exercises in post-stroke dysarthria intervention: a randomized feasibility trial. International Journal of Language & Communication Disorders. 2014 Sep-Oct;49(5):602-617.

26S. Nielsen M, Keefe FJ, Bennell K, Jull GA. Physical Therapist-Delivered Cognitive-Behavioral Therapy: A Qualitative Study of Physical Therapists' Perceptions and Experiences. Phys Ther. 2014 Feb;94(2):197-209.

27S. Presseau J, Hawthorne G, Sniehotta FF, Steen N, Francis JJ, Johnston M, *et al*. Improving Diabetes care through Examining, Advising, and prescribing (IDEA): Protocol for a theory-based cluster randomised controlled trial of a multiple behaviour change intervention aimed at primary healthcare professionals. Implement Sci. 2014;9(1).

28S. Robbins LB, Pfeiffer KA, Wesolek SM, Lo Y. Process evaluation for a school-based physical activity intervention for 6th- and 7th-grade boys: Reach, dose, and fidelity. Eval Program Plann. 2014;42:21-31.

29S. Van Schijndel-Speet M, Evenhuis HM, Van Wijck R, Echteld MA. Implementation of a group-based physical activity programme for ageing adults with ID: A process evaluation. J Eval Clin Pract. 2014;20(4):401-407.

30S. Washington T, Zimmerman S, Cagle J, Reed D, Cohen L, Beeber AS, *et al*. Fidelity decision making in social and behavioral research: Alternative measures of dose and other considerations. Soc Work Res.2014;38(3):154-162.

31S. Almas A, Islam M, Jafar TH. School-based physical activity programme in preadolescent girls (9-11 years): a feasibility trial in Karachi, Pakistan. Arch Dis Child. 2013 July;98(7):515-519.

32S. Bach E, Beissner K, Murtaugh C, Trachtenberg M, Carrington Reid M. Implementing a cognitive-behavioral pain self-management program in home health care, part 2: Feasibility and acceptability cohort study. J Geriatr Phys Ther 2013;36(3):130-137.

33S Barber SE, Jackson C, Akhtar S, Bingham DD, Ainsworth H, Hewitt C, et al. "Pre-schoolers in the playground" an outdoor physical activity intervention for children aged 18 months to 4 years old: Study protocol for a pilot cluster randomised controlled trial. Trials 2013;14(1).

34S. Benzo R, Vickers K, Ernst D, Tucker S, McEvoy C, Lorig K. Development and feasibility of a self-management intervention for chronic obstructive pulmonary disease delivered with motivational interviewing strategies. J Cardiopulm Rehabil Prev. 2013;33(2):113-122.

35S Bergström H, Hagströmer M, Hagberg J, Elinder LS. A multi-component universal intervention to improve diet and physical activity among adults with intellectual disabilities in community residences: A cluster randomised controlled trial. Res Dev Disabil 2013;34(11):3847-3857.

36S. Branscum P, Sharma M, Wang LL, Wilson B, Rojas-Guyler L. A Process Evaluation of a Social Cognitive Theory-Based Childhood Obesity Prevention Intervention: The Comics for Health Program. Health Promot Pract. 2013;14(2):189-198.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

37S. Gabbay RA, Añel-Tiangco RM, Dellasega C, Mauger DT, Adelman A, Van Horn DHA. Diabetes nurse case management and motivational interviewing for change (DYNAMIC): Results of a 2-year randomized controlled pragmatic trial. J Diabetes. 2013;5(3):349-357.

38S. Goode AD, Eakin EG. Dissemination of an evidence-based telephone-delivered lifestyle intervention: Factors associated with successful implementation and evaluation. Transl Behav Med. 2013;3(4):351-356

39S. Lorencatto F, West R, Christopherson C, Michie S. Assessing fidelity of delivery of smoking cessation behavioural support in practice. Implement Sci. 2013;8(1).

40S. Mars T, Ellard D, Carnes D, Homer K, Underwood M, Taylor SJC. Fidelity in complex behaviour change interventions: A standardised approach to evaluate intervention integrity. BMJ Open. 2013;3(11).

41S. Pfeiffer KA, Saunders RP, Brown WH, Dowda M, Addy CL, Pate RR. Study of Health and Activity in Preschool Environments (SHAPES): Study protocol for a randomized trial evaluating a multi-component physical activity intervention in preschool children. BMC Public Health. 2013;13(1).

42S. Poston L, Briley AL, Barr S, Bell R, Croker H, Coxon K, *et al*. Developing a complex intervention for diet and activity behaviour change in obese pregnant women (the UPBEAT trial); Assessment of behavioural change and process evaluation in a pilot randomised controlled trial. BMC Pregnancy and Childbirth. 2013;13.

43S. Scobbie L, McLean D, Dixon D, Duncan E, Wyke S. Implementing a framework for goal setting in community based stroke rehabilitation: A process evaluation. BMC Health Serv Res. 2013;13(1).

44S. Sears K, Blair K, Iovannone R, Crosland K. Using the Prevent-Teach-Reinforce Model with Families of Young Children with ASD. J Autism Dev Disord. 2013 May;43(5):1005-1016.

45S. Seo D, King MH, Kim N, Sovinski D, Meade R, Lederer AM. Predictors for moderate- and vigorous-intensity physical activity during an 18-month coordinated school health intervention. Prev Med. 2013;57(5):466-470.

46S. Sternfeld B, LaCroix A, Caan BJ, Dunn AL, Newton KM, Reed SD, *et al*. Design and methods of a multi-site, multi-behavioral treatment trial for menopausal symptoms: The MsFLASH experience. Contemporary Clinical Trials. 2013;35(1):25-34.

47S. Willner P, Rose J, Jahoda A, Stenfert Kroese B, Felce D, Macmahon P, *et al*. A cluster randomised controlled trial of a manualised cognitive behavioural anger management intervention delivered by supervised lay therapists to people with intellectual disabilities. Health Technol Assess. 2013;17(32):1-173

48S. Zheng H, Rosal MC, Oatis CA, Li W, Franklin PD. Tailored system to deliver behavioral intervention and manage data in randomized trials. J Med Internet Res. 2013;04;15(4)58-58.

49S. Bodde AE, Seo D, Frey GC, Lohrmann DK, Van Puymbroeck M. Developing a physical activity education curriculum for adults with intellectual disabilities. Health Promotion Practice. 2012;13(1):116-123.

50S. Broekhuizen K, Van Poppel MNM, Koppes LL, Kindt I, Brug J, Van Mechelen W. Can Multiple Lifestyle Behaviours be Improved in People with Familial Hypercholesterolemia? Results of a Parallel Randomised Controlled Trial. PLoS ONE. 2012;7(12).

51S.Brookman-Frazee LI, Drahota A, Stadnick N. Training community mental health therapists to deliver a package of evidence-based practice strategies for school-age children with autism spectrum disorders: A pilot study. J Autism Dev Disord. 2012;42(8):1651-1661.

52S. Cate H, Bhattacharya D, Clark A, Fordham R, Notley C, Broadway DC. Protocol for a randomised controlled trial to estimate the effects and costs of a patient centred educational intervention in glaucoma management. BMC Ophthalmol. 2012;12(1).

53S. Cowan JA, Devine CM. Process evaluation of an environmental and educational nutrition intervention in residential drug-treatment facilities. Public Health Nutr. 2012;15(7):1159-1167.

54S. Faulkner MS. Intervention fidelity: Ensuring application to practice for youth and families. Journal for Specialists in Pediatric Nursing. 2012;17(1):33-40.

55S. Galanter R, Self-Brown S, Valente JR, Dorsey S, Whitaker DJ, Bertuglia-Haley M, *et al*. Effectiveness of Parent–Child Interaction Therapy Delivered to At-Risk Families in the Home Setting. Child Fam Behav Ther. 2012 Jul;34(3):177-196.

56S. Heideman WH, de Wit M, Middelkoop BJC, Nierkens V, Stronks K, Verhoeff AP, et al. DiAlert: A prevention program for overweight first degree relatives of type 2 diabetes patients: Results of a pilot study to test feasibility and acceptability. Trials. 2012;13

57S. Hildebrand MW, Host HH, Binder EF, Carpenter B, Freedland KE, Morrow-Howell N, *et al*. Measuring treatment fidelity in a rehabilitation intervention study. Am J Phys Med Rehabil. 2012;91(8):715-724.

58S. Hollands GJ, Whitwell SCL, Parker RA, Prescott NJ, Forbes A, Sanderson J, *et al*. Effect of communicating DNA based risk assessments for Crohn's disease on smoking cessation: Randomised controlled trial. BMJ. 2012;345(7870).

59S. Irvine L, Falconer DW, Jones C, Ricketts IW, Williams B, Crombie IK. Can Text Messages Reach the Parts Other Process Measures Cannot Reach: An Evaluation of a Behavior Change Intervention Delivered by Mobile Phone? PLoS ONE. 2012;7(12).

60S. Knowlden A, Sharma M. A feasibility and efficacy randomized controlled trial of an online preventative program for childhood obesity: Protocol for the EMPOWER intervention. J Med Internet Res. 2012;14(3).

61S. Llewellyn C, Abraham C, Miners A, Smith H, Pollard A, Benn P, *et al*. Multicentre RCT and economic evaluation of a psychological intervention together with a leaflet to reduce risk behaviour amongst men who have sex with men (MSM) prescribed post-exposure prophylaxis for HIV following sexual exposure (PEPSE): A protocol. BMC Infect Dis. 2012;12.

E Supplement table: Characteristics of reviewed papers and summary of fidelity methods

62S. McCurry SM, Lafazia DM, Pike KC, Logsdon RG, Teri L. Development and evaluation of a sleep education program for older adults with dementia living in adult family homes. Am J Geriatr Psychiatry. 2012 Jun;20(6):494-504.

63S. Moore GF, Moore L, Murphy S. Integration of motivational interviewing into practice in the national exercise referral scheme in wales: A mixed methods study. Behav Cognitive Psychother. 2012;40(3):313-330.

64S. Morgenstern J, Kuerbis A, Amrhein P, Hail L, Lynch K, McKay JR. Motivational interviewing: A pilot test of active ingredients and mechanisms of change. Psychol Addict Behav. 2012;26(4):859-869.

65S. Robbins LB, Pfeiffer KA, Maier KS, LaDrig SM, Berg-Smith S. Treatment Fidelity of Motivational Interviewing Delivered by a School Nurse to Increase Girls' Physical Activity. J Sch Nurs. 2012; 28(1):70-78.

Table 1 National Institure of Health (NIH) Behaviour Change Consortium (BCC). Domains of Treatment Fidelity. Bellg *et al.* [1]

Design of study: Treatment fidelity practices related to study design ensure that a study adequately tests its hypotheses in relation to its underlying theoretical and clinical processes. *Training providers*: Treatment fidelity involves assessing and improving the training of treatment providers to ensure that they have been satisfactorily trained to deliver the intervention to study participants. *Delivery of treatment:* Treatment fidelity processes that monitor and improve delivery of the intervention so that it is delivered as intended *Receipt of treatment:* Receipt of treatment involves processes that monitor and improve the ability of patients to understand and perform treatment-related behavioural skills and cognitive strategies during treatment delivery. *Enactment of treatment skills:* Enactment of treatment skills consists of processes to monitor and improve the ability of patients to perform treatment-related behavioural skills and cognitive strategies in relevant real-life settings. *Definition:* Treatment fidelity refers to the methodological strategies used to monitor and enhance the reliability and validity of behavioural interventions. It also refers to the methodological practices used to ensure that a research study reliably and validly tests a clinical intervention.

Table 2. Summary of results

Reference	Definition	Study Design	Training providers	Delivery	Receipt	Enactment	Number of components
Bailey et. al 2015[1S]	No definition			✓			1/5
Beck et. al 2015[2S]	Yes (reference)[7]	✓	✓	✓	✓		4/5
Casey et. al 2015[3S]	No definition			✓			1/5
Chesworth et. al 2015[4S]	Yes (reference)[1]			✓			1/5:
Fortington et. al 2014[5S]	Yes (referenced)[13,14]					✓	1/5
French et. al 2015[6S]	Yes (reference			✓			1/5
Fulkerson et. al 2015[7S]	No definition			✓			1/5
Hanbury et. al 2015[8S]	Yes (no reference)			✓			1/5
Lawton et. al 2015[9S]	Yes (reference)[1,15,16]			✓	✓	✓	3/5
Martin et. al 2015[10S]	No definition			✓			1/5
McNamara et. al 2015[11S]	Yes (reference)[16]	✓		✓	✓		3/5
Pawar et. al 2015[12S]	No definition			✓			1/5
Pincus et al. 2015[13S]	No definition						1/5
Williams et. al 2015[14S]	No definition		✓				1/5
Winnett et. al 2015[15S]	No definition	✓	✓	✓	✓	✓	5/5
Wyatt et. al 2015[16S]	Yes (reference)[9,17,18]	✓	✓	✓	✓		4/5
Avery et. al 2014[17S]	Yes (reference)[3]		✓	✓			2/5
Baquero et. al 2014[18S]	Yes (no reference)		✓	✓		✓	3/5
Bryant et. al 2014[19S]	Yes (reference)[1]		✓	✓			2/5
Dewing et. al 2014[20S]	No definition			✓			1/5
Dyas et. al 2014[21S]	Yes (reference)[19]			✓	✓	✓	3/5
Hardeman et. al 2014[22S]	Yes (reference)[1]		✓	✓			2/5

Table 2. Summary of results

Kulwa et. al 2014[23S]	Yes (no reference)		✓	✓			2/5
Lorencatto et al 2014[24S]	Yes (reference)[20]			✓			1/5
McKenzie et. al 2014[25S]	No definition			✓			1/5
Neilson et. al 2014[26S]	No definition		✓	✓			2/5
Presseau et. al 2014[27S]	Yes (no reference)			✓	✓	✓	3/5
Robbins et. al 2014[28S]	Yes (reference)[21]			✓			1/5
Van Schijindel-Speet et. al 2014[29S]	Yes (reference)[22-24]			✓			1/5
Washington et. al 2014[30S]	Yes (reference)[25]	✓			✓	✓	3/5
Almas et. al 2013[31S]	Yes (no reference)			✓			1/5
Bach et al. 2013[32S]	No definition			✓			1/5
Barber et al. 2013[33S]	No definition		✓	✓			2/5
Benzo et. al 2013[34S]	No definition	✓	✓	✓			3/5
Bergstrom et. al 2013[35S]	Yes (reference)[26]			✓			1/5
Branscum et. al 2013[36S]	Yes (no reference).			✓			1/5
Gabbay et. al 2013[37S]	No definition		✓	✓			2/5
Goode et. al 2013[38S]	Yes (reference)[25]			✓	✓		2/5
Lorencatto et. al 2013[39S]	Yes (reference)[1,20]			✓			1/5
Mars et. al 2013[40S]	Yes (reference)[1]			✓			1/5
Pfeiffer et. al 2013[41S]	No definition			✓			1/5
Poston et. al 2013[42S]	Yes (no reference)			✓			1/5
Scobbie et. al 2013[43S]	No definition/			✓			1/5
Sears et. al 2013[44S]	No definition		✓	✓			2/5
Seo et. al	No definition			✓			1/5

Table 2. Summary of results

2013[45S]							
Sternfield et. al 2013[46S]	No definition		✓	✓			2/5
Wilner et. al 2013[47S]	Yes (reference)[2]			✓			1/4
Zheng et. al 2013[48S]	No definition			✓			1/5
Bodde et. al 2012[49S]	No definition			✓			1/5
Broekhuizen et. al 2012[50S]	No definition			✓			1/5
Brookman-Frazee et. al 2012[51S]	No definition			✓			1/5
Cate et. al 2012[52S]	No definition		✓				1/5
Cowan and Devine 2012[53S]	No definition					✓	1/5
Faulkner et. al 2012[54S]	Yes (reference)[1]	✓	✓	✓	✓	✓	5/5
Gallanter et. al 2012[55S]	No definition			✓			1/5
Heideman et. al 2012[56S]	Yes (no reference)			✓			1/5
Hildebrand et. al 2012[57S]	Yes (reference)[27,28]			✓			1/5
Hollands et. al 2012[58S]	No definition			✓			1/5
Irvine et. al 2012[59S]	Yes (no reference).			✓			1/5
Knowlden and Sharma 2012[60S]	Yes (no reference)				✓		1/5
Llewellyn et. al 2012[61S]	Yes (no reference).	✓	✓	✓			3/5
McCurry et. al 2012[62S]	No definition			✓	✓	✓	3/5
Moore et. al 2012[63S]	No definition		✓	✓			2/5
Morganstrern et. al 2012[64S]	No definition		✓	✓			2/5
Robbins et. al 2012[65S]	Yes (reference)[1]	✓	✓	✓	✓	✓	4/5